



International Conference on Communication Technology and System Design 2011

Novel Approach of Adaptive Thresholding Technique for Edge Detection in Videos

Debabrata Samanta^a, Goutam Sanyal^a, a*^{a,b}*Department of CSE, National Institute of Technology, Durgapur,
Mahatma Gandhi Avenue, West Bengal, India - 713209*

Abstract

Edge detection and motion estimation is one of the important aspect to analyzed the video .In this paper an adaptive threshold generation technique is used to filter the edges from its background from a video .We first propose an algorithm for detecting edges within video frames directly on the JPEG format without decompression process. When a vehicle crosses a speed limit, this can be highlighted by counting the number of image frames generated in a defined background region.

© 2011 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of ICCTSD 2011

Open access under [CC BY-NC-ND license](#).

Keywords: Gradient; Adaptive; Threshold; Motion Generatio; Image frame.

1. Introduction

The amount of digital video available has increased dramatically in the last few years. Video is a rich and most convenient way to get information due to advanced and friendly multimedia available techniques. Edge detection is a very important low-level image processing operation, which is essential in order to carry out various higher level tasks such as motion and feature analysis, understanding, recognition and retrieval from databases.

First a video (.avi file) is translated into image frames and then the image frames are converted into JPEG format. Then our proposed edge detection technique is applied on each of the image frames to filter the edges. All edge images are again translated into frame image. Lastly the resultant frames are translated into video file (.avi).

Application of derivative operators on intensity image produces another image, usually called gradient image as it reveals the rate of intensity variation. This image is made to undergo thresholding and/or edge linking in order to yield contours. Thus the image is decomposed into various regions resulting in another kind of segmentation .

2. Related Work

Kavitha Ganesan and Shanmugam Jalla [1] have proposed algorithms for automatic segmentation of objects in image sequences. The number of objects in the images increases, boundary inaccuracy increases in case of change detection methods. Model matching techniques partially overcomes the drawbacks of change detection method. Michael Lee, Surya Nepal, Uma Srinivasan [2] used spatial-domain synthetic edge model, which is defined

* Debabrata Samanta. Tel.: +91-9239537363.

E-mail address: debabrata.samanta369@gmail.com.

using interrelationship of two DCT edge feature, a) horizontal, b) vertical. A major drawback is that the detected edges have poor connectivity with each other because each of them is individually processed within a block boundary. Shwu-Huey Yen, Hsiao-Wei Chang, Chia-Jen Wang, Chun-Wei Wang [3] have proposed the temporal information of video and logical AND operation to remove most of irrelevant background and it processes 4 frames if the video is played on a frame rate of 4 frames per second.

Trung Quy Phan, Palaiahnakote Shivakumara and Chew Lim Tan [4] propose an efficient text detection method based on the Laplacian operator. Wei Zhao, Zuying Luo, Je-rey Fan, Sheldon X.-D. Tan [5] proposed "Vector Edge" of the current frame to store all the vectors' information within a frame and put them in the Laplacian of Gaussian edge detection operator.

3. Proposed Method

3.1 Get the frame images from video

We know that a viewer needs 2 seconds or more to process a complex scene. So the videos are played f frames per second, we are interested in detecting video object edge staying on different location for at least $2f$ consecutive frames. Now we get these frame-images and apply our proposed Adaptive Thresholding Technique for Edge Detection in Frame Images using mentioned in subsection.

3.2 Convolution with five order mask

We have proposed a 5 X 5 convolution mask. One estimating the gradient in the X direction, the other estimating the gradient in the Y direction. In this mask higher weights are assigned to the pixels close to the candidate pixel. So the gradient of the pixels is calculated as:

$$G_x(i,j) = I(i,j) * M_x = \frac{1}{5} \sum_{K=-5/2}^{5/2} \sum_{H=-5/2}^{5/2} M_x(i,j) I(i-H,j-K)$$

And other hand,

$$G_y(i,j) = I(i,j) * M_y = \frac{1}{5} \sum_{K=-5/2}^{5/2} \sum_{H=-5/2}^{5/2} M_y(i,j) I(i-H,j-K)$$

Where * indicates a discrete convolution, M_x and M_y are the 5X5 mask, and $I(i,j)$ is a $h \times w$ image. Compute the gradient magnitude approximation at each pixel (i,j) as :

$$G_{i,j} = \sqrt{G_x(i,j)^2 + G_y(i,j)^2}$$

3.3 Adaptive Thresholding

Image thresholding is a segmentation technique because it classifies pixels into two categories. Category1: Pixels whose gray level values fall below the threshold and Category2: Pixels whose gray level values are equal or exceed the threshold. In gray level image, range of input data set is $[0, 1]$. After thresholding, output data set contains only two values 0 and 1. Thus, thresholding creates a binary image. If T is a threshold value, then any pixel (x, y) for which $f(x, y) > T$ is called an object point; otherwise the pixel is called a background pixel.

For thresholding we compute adaptive threshold of local intensity variations as:

1. First the overall mean value of the gradient image is calculated. So the pixels having lower edge strength than this mean value are already discarded.

Threshold1 = mean (G)

2. Then a 3 X 3 window is splits over the gradient image where the mean and variance of the gradient image within this window are calculated. Then taking the sum of this mean value and standard deviation and this is considered as the threshold value of that pixel.

Now if the gradient of this image exceeds this threshold then the pixel is treated as edge.

If G is a gradient image then the threshold value is generated as follows:

First mean of the overall gradient image is calculated:

$$T_1 = \frac{1}{9} \sum_{K=-1}^1 \sum_{H=-1}^1 G(i+H, j+K)$$

After this the pixels which have less gradient of this mean T_1 are already discarded. Next the standard deviation of the intensity variation of the local pixels is calculated:

$$\delta = \sqrt{\frac{1}{9} \sum_{K=-1}^1 \sum_{H=-1}^1 G(i+H, j+K) - T_1}$$

Lastly this mean and standard deviation is summing up to generate the ultimate threshold for the candidate pixel $T = T_1 + \delta$.

In this way the threshold value is generated both dynamically and region-wise for each pixels so that the possibility of data loss or noise is quite reduced.

If the gradient of the pixel is greater than or equal to this adaptive threshold then only the pixel is treated as edge otherwise it is discarded.

$$G(i,j) = 1 \text{ if } G(i,j) \geq T \\ = 0 \text{ otherwise}$$

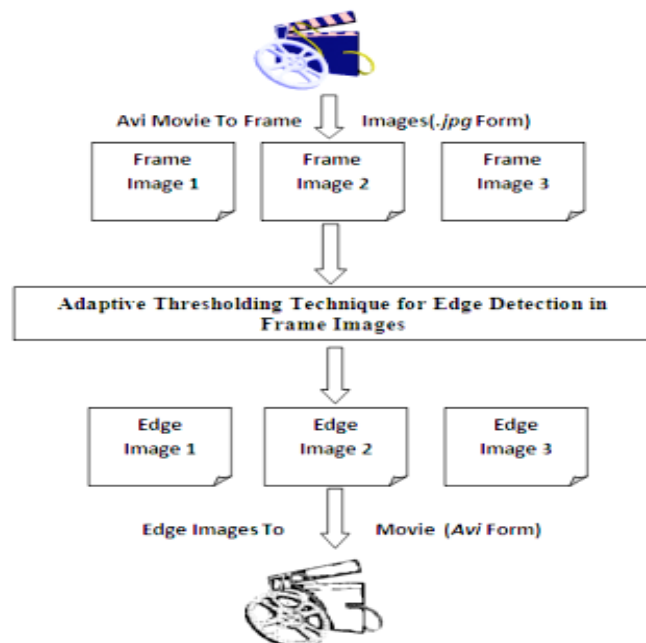
3.4 Non-Maximal Suppression (NMS)

The edges obtained now has to be sharpen and thin for this reason Canny pointed out that there should only be one response of an edge detection system to a single edge in the image. The NMS operation considers the fact that an edge at a pixel is legitimate only when the gradient magnitude at that pixel assumes a maximum in the gradient direction within a local surrounding. We calculate the gradient direction at each location in the image under consideration using the following expression:

$$\Theta = \arctan(G_Y / G_X)$$

For simplicity, the values of the directions obtained are then approximated $\Theta_{x,y}$ to the closest among the following set, $[0, 45, 90, 135]$. We then retain only those $G_{x,y}$ which are greater than the other gradient values in local surrounding and in the corresponding gradient directions $\Theta_{x,y}$.

3.5 PROPOSED DIAGRAM OF METROLOG



3.6 Proposed Algorithm

1. Start.
2. A video (.avi file) is translated into image frames and then the image frames are converted into JPEG form.
3. Read the frame image.
4. Obtain gradient of each pixel by convolution using 5X5 mask.
5. Generating threshold adaptively as follows:

If $G(x, y) \geq \text{Threshold1}$ then
begin

$$\text{mean}(x, y) = \frac{1}{9} \sum_{i=0}^n G_{(X+n, Y+n)}$$

$$\text{variance}(x, y) = \frac{1}{9} \sum_{i=0}^n G_{(X+n, Y+n)} - \text{mean}(x, y)$$

$$\delta(x, y) = \sqrt{\text{variance}(x, y)}$$

$$\text{Threshold}(x, y) = \text{mean}(x, y) + \delta(x, y)$$

end

Else the pixels are treated as non edge.

begin

$$\text{Edge}(x, y) = 1 \quad \text{if } G(x, y) \geq \text{Threshold}(x, y)$$

$$= 0 \quad \text{otherwise}$$

end

6. Suppressing the binary image using canny non maximal suppression.
7. Edge image is obtained.
8. All edge images are again translated into frame image
9. The resultant frames are translated into video file (.avi).
10. Stop.

4. Experimental Results

First, we taken a video (Fig 1) and get the frame image (Fig 2 and Fig 3). Then using our methodology we get the edge images (Fig 4 and Fig 5). Lastly we get the edge video (Fig 6).

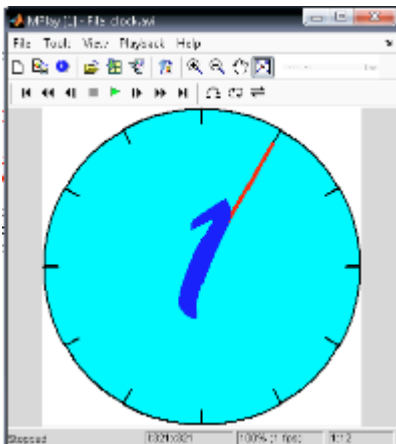


Fig 1: Input video

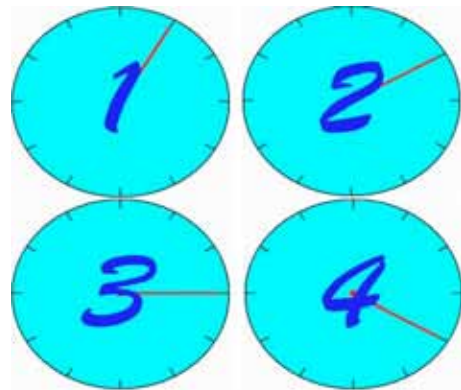


Fig 2: The frames images:

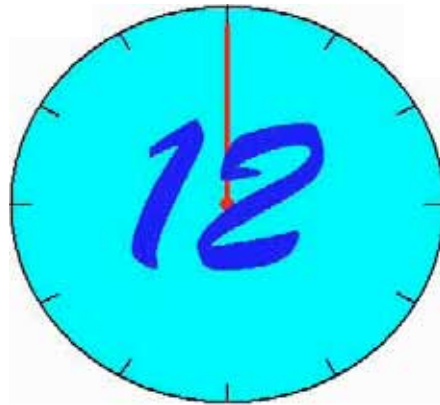


Fig 3: Last frame images

Now we get the frame edge images, and final edge image of given image:

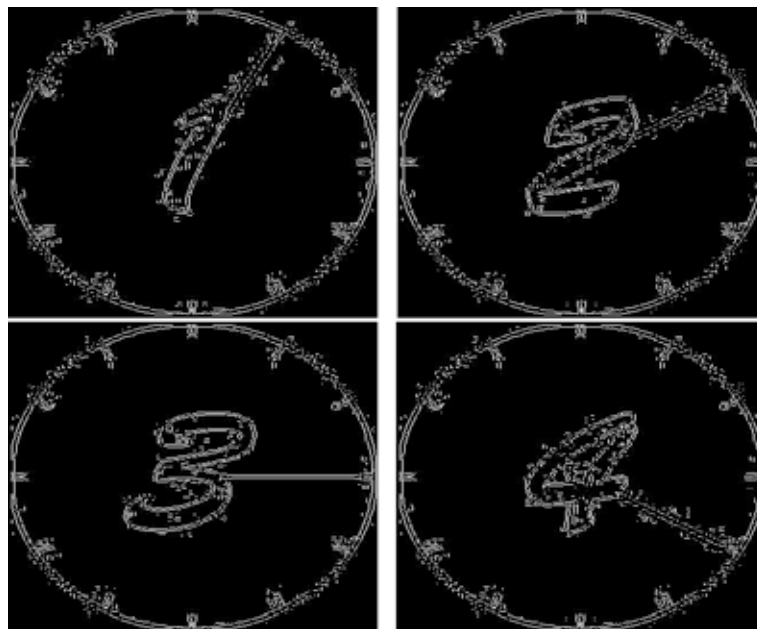


Fig 4: 1st, 2nd, 3rd and 4th frame Edge images

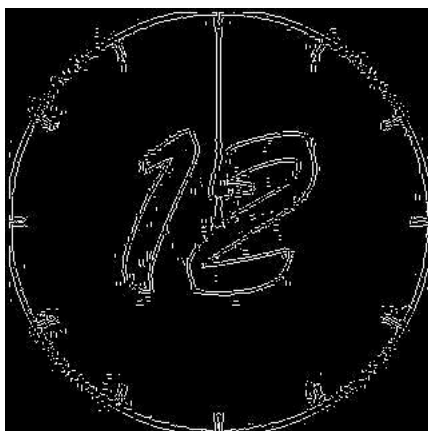


Fig 5: Last frame Edge images



Fig 6: Final Edge Video

5. Conclusion

Detecting the edges in an image using gradient value based methodologies is heavily dependent on a decision process that uses threshold values determined from the sum of gradient mean and standard deviation of the image within a 3 x 3 window.

The novel methodology described in this paper, which determines a adaptive threshold from a gradient mean and standard deviation to be used as the threshold of that pixel, thus the threshold value is generated dynamically for different regions as well as no manual threshold is required to provide. And applying the edge detection on video gives a better way to generate motion information of an moving object. In general traffic rules there exists a speed boundary. If any car crosses this speed limit then it breaks the traffic rule. So the frame image generated from the video of the moving object within that given speed limit is a certain number. Therefore if any car crosses this speed limit, the number of frame images generated from the video of that moving car exceeds that given number. Thus the car may be identified for breaking the traffic rule.

6. References

- [1] Kavitha Ganesan and Shanmugam Jalla, " Video Object Extraction Based on a Comparative Study of Efficient Edge Detection Techniques", *Proc. The International Arab Journal of Information Technology*, April 2009, Vol. 6, No. 2.
- [2] Michael Lee, Surya Nepal, Uma Srinivasan," Role of Edge Detection in Video Semantics",*Proc. Conferences in Research and Practice in Information Technology*, 2003,Vol. 22.
- [3] Shwu-Huey Yen, Hsiao-Wei Chang, Chia-Jen Wang, Chun-Wei Wang, Proc, " WSEAS TRANSACTIONS on SIGNAL PROCESSING", October 2010, Issue 4, Volume 6,.
- [4] Trung Quy Phan, Palaiahnakote Shivakumara and Chew Lim Tan,"A Laplacian Method for Video Text Detection", *10th International Conference on Document Analysis and Recognition*, 2009, IEEE DOI 10.1109/ICDAR.2009.153.
- [5] Wei Zhao, Zuying Luo, Je-rey Fan, Sheldon X.-D," Vector Edge Detection in H.264 Implementation", *The 2008 International Conference on Embedded Software and Systems Symposia (ICESS2008)*, 2008, IEEE DOI 10.1109/ICESS. Symposia.2008.31,.
- [6] Marimont, D.H., Rubner, Y., "A probabilistic framework for edge detection and scale selection", *Computer Vision, 1998. Sixth International Conference on* Issue Date: 4-7 Jan 1998 On page(s): 207- 214.
- [7] Lindeberg T., "Edge Detection and Ridge Detection with Automatic Scale Selection", *International Journal of Computer Vision*, 1998, vol. 30, number 2, 117-154.
- [8] Narendra Ahuja, "A Transform for Multiscale Image Segmentation by Integrated Edge and Region Detection", *IEEE Computer Society*, Dec 1996 Volume: 18 Issue:12.
- [9] J. Jarvis and C. Roberts, "A new technique for displaying continuous tone images on a bilevel display", *IEEE Trans. Commun.*, Aug. 1976, pp. 891-898.
- [10] H. Wong and L. Guan, "A Neural Learning Approach for Adaptive Image Restoration Using a Fuzzy Model-based Network Architecture", *IEEE Trans*, 2001, Neural Network 12 (3) 516-531,.
- [11] M. Tabb and N. Ahuja, "Multiscale image segmentation using a recent transformation", in Proc. ARPA Image Understand. Workshop, 1994, vol.2.12. B. Widrow and S. Stearns, Adaptive Signal Processing. Englewood Cliffs, NJ: Prentice-Hall, 1985.
- [12] Venkatesh, Svetha, Rosin, Paul L., "Dynamic-threshold determination by local and global edge evaluation", *Applications of Artificial Intelligence* 1993, Vol. 1964, p. 40-50.
- [13] Marimont, D., Rubner Y., "A Probabilistic Framework for Edge Detection and Scale Selection", *Proceedings of the IEEE International Conference on Computer Vision*, January 1998, 207-214.
- [14] Lindeberg T., "Edge Detection and Ridge Detection with Automatic Scale Selection", *International Journal of Computer Vision*, 1998, vol. 30, number 2, 117-154,.